E7.3 10514
CR-131472

PROGRESS REPORT

February 20, 1973 - April 19, 1973

Crop Identification and Acreage Measurement Utilizing ERTS Imagery 013

Principle Investigator

Donald H. Von Steen AG 328

N73-22288

(E73-10514) CROP IDENTIFICATION AND ACREAGE MEASUREMENT UTILIZING ERTS IMAGERY Progress Report, 20 Feb. - 19 Apr. 1973 (Department of Agriculture) 5 p HC $3.00 CSCL 02C G3/13 Unclas 00514

Data Acquisition

Satellite Imagery

The Penn State Classifier schemes, Normal Map and Uniform Map, are being used to printout ERTS bulk data tapes. Uniform Map computes and prints out the difference between adjacent points. Normal Map measures the intensity of each point and prints out different symbols for different intensities.

Segments have been located on Normal Map printouts of areas in Kansas. The printouts for Kansas are most encouraging because of the large field sizes which make location of "training areas" for the classifiers easier to determine. These segments are being subsetted for classification. The Normal Map printouts are being used to locate the segments. However, it is expected that Uniform Map will be useful in later work.

Aerial Photography

Some of the RC-8 coverage of Kansas, Missouri, and South Dakota for August 18, 1972, August 28, 1972, and September 22, 1972 respectively, was sent to Photometric Data Systems. They are scanning the designated areas with their Microdensitometer. This data will allow us to work on developing data handling techniques, locate fields, and practice classifying before the Microdensitometer is received and in-house data is available.

In regard to Mr. Freden's letter on January 31, 1973:

1) The aircraft support provided has generally been adequate. Aircraft support provided coverage on two and in some cases three dates. Due to the lateness of satellite launching, additional support was not deemed necessary.

2) Since the acquisition of a microdensitometer has been delayed, time between overflight and receipt of photography has not been a problem in this investigation. However, at least a month expired between overflights and receipt of photography and this would be a serious delay in an operational survey.

3) The aircraft data is being used to aid in interpretation of satellite imagery and for quality control of ground enumeration. It will be digitized on a microdensitometer and this data will be used to supplement the satellite imagery and ground observations by supplying missing data. Clouds prevent complete coverage with satellite imagery and respondent refusals are a limitation of ground enumeration. Assuming we are able to accurately discriminate and classify the digitized data, we will attempt to combine the results with ground data and satellite imagery.

Microdensitometer Purhcase

An invitation for bid for a Microdensitometer Data Acquisition System, including a Computer Control System, was prepared and submitted to GSA for review and delegated procurement authority prior (DPA) to submission to industry. The Solicitation and requested DPA were returned and approved by GSA. Bids were opened on March 15, 1973. Only two bids were received. Both bids were non responsive. One failed to comply to the terms of the solicitation and the other firm qualified its bid with respect to payments. However, the latter firm met the technical specifications of the solicitation. Since they did meet the technical specifications, we have asked GSA to permit us to negotiate with them under a new delegation of procurement authority. We are now awaiting GSA's approval.

Data Processing Software

Segment Location

Computer programs to compute the location and to extract the data for individual segments within any given set of ERTS MSS data tapes have been written and tested. Some problems which still exist at this point are:

1) The digital data contained by the MSS system corrected tapes for picture 1061-16570 does not correspond to the area shown by the 70mm transparencies. Approximately the first 120-126 scan lines on the digital tapes for picture 1061-16570 are for the area to the north of the picture, that is, for the area on picture 1061-16564. Also, approximately 42 scan lines which are needed to complete the digital tape coverage of picture 1061-16570 on the south are not on the tape. This gives a net surplus of about 84 scan lines over and above that needed to cover the area in the transparency.

2) The area covered by the transparency is supposed to be 96.3 by 100.0 nautical miles. However, 2.340 scan lines times 79.0 meters (the specified distance between scan lines) results in a digital data picture which is 99.8 nautical miles along the major axis. rather than the specified 96.3 nautical miles.

The problems presented by the above findings are:

1) Where on the digital tape is the indicated format center, and

2) If we find the location of the format center on one tape, will it be in the same place on another picture?

Other, lesser, problems include deriving a procedure for obtaining the mean distances between picture elements in a scan line and between scan lines.

We have obtained gray-scale maps of digital data, both from the Penn State system and from programs developed in-house. We have been able to identify some individual segments on these computer printouts. Hopefully, this ability will enable us to solve the remaining problems.

Systems Status (PENN STATE CLASSIFIER)

Four of the 7 main programs are now up and running. These are PSUBSET, PNMAP, PUMAP, PTPINFO.

PSUBSET is used to subset a large file into smaller rectangular areas or blocks. Each block is defined by line and element number. ERTS bulk tapes and subset tapes can be subsetted by PSUBSET.

PNMAP assigns mapping symbols to all points with given contrast level. It is used to give line printer gray-scale maps.

PUMAP assigns mapping symbols based on the contrast difference. It is used to accentuate field boundaries.

PTPINFO is used to print the heading and table of contents records of a standard system file.

The reread problem discussed in the previous report was solved by substitution. The original REREAD was replaced with @INCORE. The @INCORE routine has two advantages over REREAD, (1) fewer system dependencies and, (2) provides both formatted reads and writes in core.

The remaining 3 routines PACLASS, PACLUS, and PSTATS are not running because of duplicate support subroutine names for different subroutines. Our library requires that all subroutine names be unique. We have to go back into the original program file and manually extract the required subroutines and give them a different name before they can be added to our library. This operation is underway.

Data Analysis

Since we plan to sample points prior to aerial photography classification, an optimum sample allocation has been studied. Preliminary results indicate the optimum sample allocation is about twenty scan lines or scan columns, and about twenty-five points within each scan column. This indicates that we need about 500 data points from each field. We feel this is an upper bound and probably fewer points within a field will be used in classifying.

Preliminary computer runs with our parametric discriminant functions program have been made comparing different a priori probabilities. Correct a priori probabilities do not seem to drastically improve classification on a field-by-field basis, but do help remove bias introduced without use of a priori probabilities.

The following illustration will indicate how this occurs. Let us say 100 fields of equal size are to be classified. If there are actually 10 corn fields, 50 soybean fields, and 40 wheat fields, but we ignore the prior probabilities, the computer will automatically assume the groups have an equal chance of being selected (i.e. 1/3 will be corn, 1/3 will be wheat, and 1/3 will be soybeans). The results of the classification will tend in the direction of these probabilities and possibly 20 fields will be classified corn, 45 fields soybeans, and 35 fields classified as wheat. Errors of misclassification will not compensate from crop to crop.

If we had "guessed" that corn, soybeans, and wheat had been in proportion of 2:4:4, the results may have indicated 15 fields corn, 50 fields soybeans, and 35 wheat. The misclassification errors, therefore, tend to compensate from crop to crop with inclusion of a priori probabilities.